

NERC ARSF

Data Management Plan

Issue No.	Date	Description of Changes
1	26 th June 2009	First Issue
1a	20 th July 2009	Removed requirement for users (of data other than scanned aerial photographs) to confirm data are for bona-fide research purposes only.

This work is in the Public Domain. To view a copy of the public domain certification, visit <http://creativecommons.org/licenses/publicdomain/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

1. Scope

The purpose of this data management plan is to set up a coherent approach to data issues for ARSF. Its objective is to ensure that:

- A high quality documented data archive is created.
- Appropriate data support is provided to ARSF data users and data providers.
- Data are made available to users in a timely fashion.
- Academic credit for data creation is given.
- Conditions of use, access and deposit are clearly stated and do not infringe on the data creators' rights.
- Potentially scientifically valuable data are kept for reuse in the long-term and by other disciplines.
- Results can be checked and validated.

This document is an agreed record of the data management needs and issues of the NERC Airborne Research and Survey Facility. It defines who is responsible for data management activities in ARSF, ARSF-DAN (Data Analysis Node) and NEODC (NERC EO Data Centre). It lists the expected data products and provides a mechanism for recording and agreeing changes. Other data needs and issues are also laid out so that problems can be identified early. It includes conditions of use and deposit to clearly express the ownership and rights associated with the data.

The scope of this document is data management of the data used or generated by the NERC ARSF.

This data management plan has been agreed between the NERC Earth Observation Data Centre and the Head of ARSF. In the event of dispute relating to the ARSF data the final decision rests with the NERC Data Management Coordinator.

2. About the dataset

2.1 *Definition*

The ARSF dataset encompasses all data produced by ARSF which is to be archived for posterity. This may include validated raw data and processed data, where they have long term importance and/or use to the wider scientific community. The individual core data sources are summarised in Appendix A.

Non-core data are data generated by user supplied instruments onboard the aircraft, which are operated by scientists conducting experiments.

2.2 *Archive location*

The ARSF archive is located at the NERC Earth Observation Data Centre (NEODC), part of the Centre of Environmental Data Archival (CEDA) at STFC (Science and Technology Facilities Council).

2.3 *Format*

In order to facilitate future data use, data must conform to agreed format standards. Details are given in Appendix A.

2.4 *Information on the data*

Metadata (i.e. information on the data) are a crucial part of any data archive since they ensure the accessibility and readability of the data. It is therefore essential that metadata be submitted at the same time as the data sets to which they pertain.

In addition to these 'internal' metadata, ARSF are encouraged to archive at NEODC all relevant information, including experiment descriptions, references, papers, reports, etc. These will be added to the CEDA document repository.

2.5 *Ownership of data*

Ownership of the data normally lies with NERC. Data generated under commercial contract is regarded as owned by the purchaser and is out of scope for data management by NEODC.

2.6 *Data archival*

All validated processed data (i.e. data sets in their final form) and raw data where appropriate will be archived at the NEODC. Archival must take place no later than one year after data acquisition. It is expected that delivery from ARSF directly to the PI is within one month of data acquisition.

ARSF are required to agree to the data deposit conditions before the data are added to the archive.

Deposit Conditions

1. The NERC Facility depositing the data confirms that NERC is the owner of the data and/or has the right to deposit the data in the NEODC ARSF archive
2. Ownership of the data remains with NERC
3. The NERC Data Centre reserves the right to store the data, and make the data available under the Conditions of Use
4. The NERC Facility grants the NERC Data Centre permission to, without changing content, translate the data to any medium or format for the purpose of future preservation and accessibility

Data will be archived at NEODC in one of two ways:

- A) processed data and documentation
- B) raw data, ancillary data and other inputs used by ARSF-DAN in data processing

Data in the NEODC-A archive should conform to agreed format and metadata standards and will be curated by NEODC for posterity and made available to registered and authorised users.

Data in the NEODC-B archive is the responsibility of ARSF-DAN and will be left in its original format (zipped tarfile). The NEODC-B archive data will only be made available to the ARSF and ARSF-DAN teams.

Some data may have been acquired but not processed (e.g. from instruments not requested by project PI). These data will be stored in the NEODC-B archive but ARSF must submit appropriate metadata if they want these data to be discoverable by users for possible future processing.

ARSF non-core data produced under the auspices of a NERC or other research programme will be subject to the data protocol of that programme, and will be archived at NEODC if appropriate.

2.7 *Data distribution*

Access to all data submitted to the NEODC will be restricted to ARSF project PI teams for one year following the data delivery date¹, after which they will be released to any registered user. During the embargo period, access may be extended to external collaborators who have been authorised by the project PI or their delegated authority. Potential users of the archive will be required to agree to the Conditions of Use.

Conditions of use

1. Access to all data submitted to the NEODC will be restricted to the ARSF project PI team for one year following the archival date, after which they will be released to any registered user.
2. No data should be transferred to a third party, unless they have registered at NEODC and agreed to the conditions of use
3. In all cases where the data are used in a presentation or publication, an acknowledgement must be given: for example, "*Data from the NERC ARSF are provided courtesy of NERC via the NERC Earth Observation Data Centre (NEODC).*"
4. Information submitted in application for access to the data will be made available to: The Natural Environment Research Council (NERC) and its delegated authorities, i.e. the NERC Earth Observation Data Centre and the British Atmospheric Data Centre (BADC) and their host organisation, the Science and Technology Facilities Council (STFC) for the purposes of tracking data usage and of improving the service.

3. NEODC responsibilities

NEODC will

- provide a data archive for ARSF data with password protected access for authorised users only
- archive relevant external metadata (flight logs, science summaries) if provided
- provide assistance with format and metadata issues
- offer a helpdesk service to handle queries from data users and providers
- add metadata to searchable catalogue and map interface
- create preview images of scanned aerial photographs

Appendix A: ARSF Data Streams

¹ Data delivery date = date of most recent delivery of data from ARSF to project PI

Appendix A – ARSF data streams

Type	Format	ARSF deliver to PI?	Archival policy	Notes
ATM				
Raw	Internal to Azimuth Systems, can only be handled with az* tools	No	NEODC B	
processed	Hdf	Yes	NEODC A	
CASI				
Raw	ITRES internal format. Handled with az* tools.	No	NEODC B	
processed	Hdf	Yes	NEODC A	
Eagle				
Raw	ENVI BIL	No (rare exceptions)	NEODC B	
processed	Hdf (metadata and navigation) + ENVI BIL (image data) ²	Yes	NEODC A	
Hawk				
Raw	ENVI BIL	No (rare exceptions)	NEODC B	
Processed	Hdf (metadata and navigation) + ENVI BIL (image data)	Yes	NEODC A	
Analogue photography				
Analogue film	9x9 negatives	No	BGS	
Scans	GEOTIFF	yes (via ARSF-DAN)	NEODC A	Digitisation organised by ARSF. Data delivered directly to NEODC from ARSF-Ops for archival.
Digital photography				
Raw	2007 camera: TIFF? Leica: unknown	No	NEODC B	
Processed	16 bit TIFF	Yes	NEODC A	Data to NEODC via ARSF-DAN
LiDAR				
Raw	Optec (<2009), Leica (2009	No	NEODC B	

² split into BIL and HDF as image data exceeds the filesize limits of HDF4

	onwards): proprietary formats			
processed	Ascii (xyz) point clouds.	Yes	NEODC A	Data provided to NEODC via ARSF-DAN. Improved format to be agreed.

A – archive accessible to PI team for one year after data delivery date, then all registered users

B – archive accessible to ARSF team only (ARSF and ARSF-DAN)

Type	Format	ARSF deliver to PI?	Archival policy	Notes
AIMMS	?	?	?	
?				
?				
Ancillary data				Used by ARSF DAN in data processing
Navigation: Applanix, Leica	SBET (Smoothed Best Estimate of Trajectory) binary	No (relevant subsections stored in processed HDF data)		
GPS base station data	RINEX	No		
DEM	Azgcrr: ascii grid with one line custom header	No		
Other				
Example screenshots	Jpg	Yes	NEODC A	
Flight log	Doc (ARSF-DAN convert to PDF)	Yes	NEODC A	

NOTES:

- Need to think about relevant associated data e.g. field data (FSF) and where/how they are archived
- Need to resolve how atmospheric project data are archived